Molecular Informatics 2017 vol.36 N11

Energy-based Neural Networks as a Tool for Harmony-based Virtual Screening

Zhokhova N., Baskin I.
Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. In Energy-Based Neural Networks (EBNNs), relationships between variables are captured by means of a scalar function conventionally called “energy”. In this article, we introduce a procedure of “harmony search”, which looks for compounds providing the lowest energies for the EBNNs trained on active compounds. It can be considered as a special kind of similarity search that takes into account regularities in the structures of active compounds. In this paper, we show that harmony search can be used for performing virtual screening. The performance of the harmony search based on two types of EBNNs, the Hopfield Networks (HNs) and the Restricted Boltzmann Machines (RBMs), was compared with the performance of the similarity search based on Tanimoto coefficient with “data fusion”. The AUC measure for ROC curves and 1 %-enrichment rates for 20 targets were used in the benchmarking. Five different scores were computed: the energy for HNs, the free energy and the reconstruction error for RBMs, the mean and the maximum values of Tanimoto coefficients. The performance of the harmony search was shown to be comparable or even superior (significantly for several targets) to the performance of the similarity search. Important advantages of using the harmony search for virtual screening are very high computational efficiency of prediction, the ability to reveal and take into account regularities in active structures, flexibility and interpretability of models, etc.

http://dx.doi.org/10.1002/minf.201700054

Keywords

harmony search, Hopfield nets, neural networks, Restricted Boltzmann Machines, virtual screening

References

[1] Y. LeCun, S. Chopra, R. Hadsell, M. A. Ranzato, F. J. Huang, in Predicting Structure Data (Eds.: G. Bakir, T. Hofmanb, B. Schoelkopf, A. J. Smola, B. Taskar, S. V. N. Vishwanathan), The MIT Press, Cambridge, Massachusetts, 2007, pp. 191–246.

[2] N. Nikolova, J. Jaworska, QSAR Comb. Sci. 2003, 22, 1006–1026;

[3] A. Bender, R. C. Glen, Org. Biomol. Chem. 2004, 2, 3204–3218;

[4] H. Eckert, J. Bajorath, Drug Discovery Today 2007, 12, 225–233;

[5] D. Stumpfe, J. Bajorath, Wiley Interdiscip. Rev.: Comput. Mol. Sci. 2011, 1, 260–282.

[6] A. Varnek, I. Baskin, J. Chem. Inf. Mod. 2012, 52, 1413–1437.

[7] D. M. J. Tax, Doctor Thesis, Technische Universiteit Delft (Delft), 2001;

[8] S. S. Khan, M. G. Madden, Lect. Notes Comput. Sci. 2010, 6206, 188–197;
[9] I. I. Baskin, N. Kireeva, A. Varnek, Mol. Inf. 2010, 29, 581–587.
[10] P. V. Karpov, I. I. Baskin, V. A. Palyulin, N. S. Zefirov, Dokl. Chem. 2011, 437, 107–111;
[11] P. V. Karpov, I. I. Baskin, N. I. Zhokhova, M. B. Nawrozkij, A. N. Zefirov, A. S. Yablokov, I. A. Novakov, N. S. Zefirov, Russ. Chem. Bull. 2011, 60, 2418–2424;
[12] P. V. Karpov, I. I. Baskin, N. I. Zhokhova, N. S. Zefirov, Dokl. Chem. 2011, 440, 263–265.
[13] P. V. Karpov, D. I. Osolodkin, I. I. Baskin, V. A. Palyulin, N. S. Zefirov, Bioorg. Med. Chem. Lett. 2011, 21, 6728–6731.
[14] J. Zupan, J. Gasteiger, Neural Networks in Chemistry, Wiley-VCH, Weinheim, 1999;
[15] N. M. Halberstam, I. I. Baskin, V. A. Palyulin, N. S. Zefirov, Russ. Chem. Rev. 2003, 72, 629–649;
[16] I. I. Baskin, V. A. Palyulin, N. S. Zefirov, Methods Mol. Biol. 2008, 458, 137–158;
[17] I. I. Baskin, D. Winkler, I. V. Tetko, Expert Opin. Drug Discovery 2016, 11, 785–795.
[18] J. J. Hopfield, Proc. Natl. Acad. Sci. USA 1982, 79, 2554–2558.
[19] J. J. Hopfield, D. W. Tank, Science 1986, 233, 625–633.
[20] J. J. Hopfield, D. W. Tank, Biol Cybern 1985, 52, 141–152.
[21] W. Li, N. M. Nasrabadi, IEEE Int. Joint Conf. Neural Networks 1989, 2, 287–290.
[22] A. Jagota, IEEE Trans. Neural Netw. 1995, 6, 724–735.
[23] J. P. Doucet, A. Panaye, SAR QSAR Environ. Res. 1998, 8, 249–272.
[24] M. Arakawa, K. Hasegawa, K. Funatsu, J. Chem. Inf. Comput. Sci. 2003, 43, 1390–1395;
[25] M. Arakawa, K. Hasegawa, K. Funatsu, J. Chem. Inf. Comput. Sci. 2003, 43, 1396–1402.
[26] Y. S. Abu-Mostafa, J. St. Jacques, IEEE Trans. Inf. Theory 1985, 31, 461–464.
[27] D. H. Ackley, G. E. Hinton, T. J. Sejnowski, Cognit. Sci. 1985, 9, 147–169;
[28] G. E. Hinton, T. J. Sejnowski, in Parallel Distributed Processing, Vol. 1, MIT Press, Cambridge, MA, 1986, pp. 282–317.
[29] G. E. Hinton, in Neural Networks: Tricks of the Trade, Springer, Berlin Heidelberg, 2012, pp. 599–619.
[30] P. Smolensky, in Parallel Distributed Processing, Vol. 1 (Eds.: D. E. Rumelhart, J. L. McClelland), MIT Press, Cambridge, 1986, pp. 194–281.
[31] G. E. Hinton, Neural Comput. 2002, 14, 1771–1800.
[32] G. E. Hinton, R. R. Salakhutdinov, Science 2006, 313, 504–507.
[33] Y. LeCun, Y. Bengio, G. Hinton, Nature 2015, 521, 436–444.
[34] R. Salakhutdinov, G. Hinton, Neural Comput. 2012, 24, 1967–2006.
[35] N. Huang, B. K. Shoichet, J. J. Irwin, J. Med. Chem. 2006, 49, 6789–6801.
[36] R. Guha, J. Stat. Software 2007, 18.
[37] C. Steinbeck, Y. Han, S. Kuhn, O. Horlacher, E. Luttmann, E. Willighagen, J. Chem. Inf. Comput. Sci. 2003, 43, 493–500.
[38] T. Sing, O. Sander, N. Beerenwinkel, T. Lengauer, Bioinformatics 2005, 21, 7881.
[39] N. H. Farhat, D. Psaltis, A. Prata, E. Paek, Appl. Optics 1985, 24, 1469–1475;
[40] J.-S. Jang, S.-W. Jung, S.-Y. Lee, S.-Y. Shin, Opt. Lett. 1988, 13, 248–250.
[41] D. Ventura, T. Martinez, in Artificial Neural Nets and Genetic Algorithms, Springer, Vienna, 1999, pp. 22–27.